

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/009,802B

DATE: 01/10/2000  
TIME: 16:31:13

Input Set: I009802B.RAW

This Raw Listing contains the General Information  
Section and up to first 5 pages.

fs

ENTERED

```

1  <110> APPLICANT: McCarthy, Sean A.
2  <120> TITLE OF INVENTION: NOVEL CRSP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
3  THEREFOR
4  <130> FILE REFERENCE: MEI-008
5  <140> CURRENT APPLICATION NUMBER: US/09/009,802B
6  <141> CURRENT FILING DATE: 1998-01-20
7  <150> EARLIER APPLICATION NUMBER: 08/842,898
8  <151> EARLIER FILING DATE: 1997-04-17
9  <150> EARLIER APPLICATION NUMBER: 60/071,589
10 <151> EARLIER FILING DATE: 1998-01-15
11 <160> NUMBER OF SEQ ID NOS: 19
12 <170> SOFTWARE: PatentIn Ver. 2.0
13 <210> SEQ ID NO 1
14 <211> LENGTH: 2479
15 <212> TYPE: DNA
16 <213> ORGANISM: Homo sapiens
17 <220> FEATURE:
18 <221> NAME/KEY: CDS
19 <222> LOCATION: (38)..(1087)
20 <220> FEATURE:
21 <223> OTHER INFORMATION: 'n' at position 1146 may be any nucleotide
22 <400> SEQUENCE: 1
23   ggcacgaggg ggcggcggct gcgggcgcag agcggag atg cag cgg ctt ggg gcc 55
24                                     Met Gln Arg Leu Gly Ala
25                                     1           5
26   acc ctg ctg tgc ctg ctg ctg gcg gcg gcg gtc ccc acg gcc ccc gcg 103
27   Thr Leu Leu Cys Leu Leu Leu Ala Ala Ala Val Pro Thr Ala Pro Ala
28               10           15           20
29   ccc gct ccg acg gcg acc tcg gct cca gtc aag ccc ggc ccg gct ctc 151
30   Pro Ala Pro Thr Ala Thr Ser Ala Pro Val Lys Pro Gly Pro Ala Leu
31               25           30           35
32   agc tac ccg cag gag gag gcc acc ctc aat gag atg ttc cgc gag gtt 199
33   Ser Tyr Pro Gln Glu Glu Ala Thr Leu Asn Glu Met Phe Arg Glu Val
34               40           45           50
35   gag gaa ctg atg gag gac acg cag cac aaa ttg cgc agc gcg gtg gaa 247
36   Glu Glu Leu Met Glu Asp Thr Gln His Lys Leu Arg Ser Ala Val Glu
37   55           60           65           70
38   gag atg gag gca gaa gaa gct gct gct aaa gca tca tca gaa gtg aac 295
39   Glu Met Glu Ala Glu Glu Ala Ala Ala Lys Ala Ser Ser Glu Val Asn
40               75           80           85
41   ctg gca aac tta cct ccc agc tat cac aat gag acc aac aca gac acg 343
42   Leu Ala Asn Leu Pro Pro Ser Tyr His Asn Glu Thr Asn Thr Asp Thr
43               90           95           100
44   aac gtt gga aat aat acc atc cat gtg cac cga gaa att cac aag ata 391

```

PAGE: 2

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/009,802B

DATE: 01/10/2000  
TIME: 16:31:13

Input Set: I009802B.RAW

```

45   Asn Val Gly Asn Asn Thr Ile His Val His Arg Glu Ile His Lys Ile
46           105                      110                      115
47   acc aac aac cag act gga caa atg gtc ttt tca gag aca gtt atc aca   439
48   Thr Asn Asn Gln Thr Gly Gln Met Val Phe Ser Glu Thr Val Ile Thr
49           120                      125                      130
50   tct gtg gga gac gaa gaa ggc aga agg agc cac gag tgc atc atc gac   487
51   Ser Val Gly Asp Glu Glu Gly Arg Arg Ser His Glu Cys Ile Ile Asp
52   135                      140                      145                      150
53   gag gac tgt ggg ccc agc atg tac tgc cag ttt gcc agc ttc cag tac   535
54   Glu Asp Cys Gly Pro Ser Met Tyr Cys Gln Phe Ala Ser Phe Gln Tyr
55           155                      160                      165
56   acc tgc cag cca tgc cgg ggc cag agg atg ctc tgc acc cgg gac agt   583
57   Thr Cys Gln Pro Cys Arg Gly Gln Arg Met Leu Cys Thr Arg Asp Ser
58           170                      175                      180
59   gag tgc tgt gga gac cag ctg tgt gtc tgg ggt cac tgc acc aaa atg   631
60   Glu Cys Cys Gly Asp Gln Leu Cys Val Trp Gly His Cys Thr Lys Met
61           185                      190                      195
62   gcc acc agg ggc agc aat ggg acc atc tgt gac aac cag agg gac tgc   679
63   Ala Thr Arg Gly Ser Asn Gly Thr Ile Cys Asp Asn Gln Arg Asp Cys
64           200                      205                      210
65   cag ccg ggg ctg tgc tgt gcc ttc cag aga ggc ctg ctg ttc cct gtg   727
66   Gln Pro Gly Leu Cys Cys Ala Phe Gln Arg Gly Leu Leu Phe Pro Val
67   215                      220                      225                      230
68   tgc aca ccc ctg ccc gtg gag ggc gag ctt tgc cat gac ccc gcc agc   775
69   Cys Thr Pro Leu Pro Val Glu Gly Glu Leu Cys His Asp Pro Ala Ser
70           235                      240                      245
71   cgg ctt ctg gac ctc atc acc tgg gag cta gag cct gat gga gcc ttg   823
72   Arg Leu Leu Asp Leu Ile Thr Trp Glu Leu Glu Pro Asp Gly Ala Leu
73           250                      255                      260
74   gac cga tgc cct tgt gcc agt ggc ctc ctc tgc cag ccc cac agc cac   871
75   Asp Arg Cys Pro Cys Ala Ser Gly Leu Leu Cys Gln Pro His Ser His
76           265                      270                      275
77   agc ctg gtg tat gtg tgc aag ccg acc ttc gtg ggg agc cgt gac caa   919
78   Ser Leu Val Tyr Val Cys Lys Pro Thr Phe Val Gly Ser Arg Asp Gln
79           280                      285                      290
80   gat ggg gag atc ctg ctg ccc aga gag gtc ccc gat gag tat gaa gtt   967
81   Asp Gly Glu Ile Leu Leu Pro Arg Glu Val Pro Asp Glu Tyr Glu Val
82   295                      300                      305                      310
83   ggc agc ttc atg gag gag gtg cgc cag gag ctg gag gac ctg gag agg   1015
84   Gly Ser Phe Met Glu Glu Val Arg Gln Glu Leu Glu Asp Leu Glu Arg
85           315                      320                      325
86   agc ctg act gaa gag atg gcg ctg agg gag cct gcg gct gcc gcc gct   1063
87   Ser Leu Thr Glu Glu Met Ala Leu Arg Glu Pro Ala Ala Ala Ala Ala
88           330                      335                      340
89   gca ctg ctg gga agg gaa gag att tagatctgga ccaggctgtg ggtagatgtg   1117
90   Ala Leu Leu Gly Arg Glu Glu Ile
91           345                      350
W--> 92   caatagaaat agctaattta tttccccang_tgtgtgcttt aagcgtgggc tgaccaggct 1177
93   tcttcttaca tcttcttccc agtaagtttc cctctgtggt tgacagcatg aggtgttgtg 1237
94   catttgttca gctccccag gctgttctcc aggcttcaca gtctggtgct tgggagagtc 1297

```

PAGE: 3

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/009,802B

DATE: 01/10/2000  
TIME: 16:31:13

Input Set: I009802B.RAW

```

95      aggcagggtt aaactgcagg agcagtttgc caccctgtc cagattattg gctgctttgc 1357
96      ctctaccagt tggcagacag ccgtttgttc tacatggctt tgataattgt ttgaggggag 1417
97      gagatggaaa caatgtggag tctccctctg attggttttg gggaaatgtg gagaagagtg 1477
98      ccctgctttg caaacatcaa cctggcaaaa atgcaacaaa tgaattttcc acgcagttct 1537
99      ttccatgggc ataggtaagc tgtgccttca gctgttgagc atgaaatgtt ctgttcaccc 1597
100     tgcattacat gtgtttattc atccagcagt gttgctcagc tcctacctct gtgccagggc 1657
101     agcattttca tatccaagat caattccctc tctcagcaca gcctggggag ggggtcattg 1717
102     ttctcctcgt ccatcaggga tttcagaggc tcagagactg caagctgctt gcccaagtca 1777
103     cacagctagt gaagaccaga gcagtttcat ctggttgtga ctctaagctc agtgctctct 1837
104     ccactacccc acaccagcct tggtgccacc aaaagtgtc cccaaaagga aggagaatgg 1897
105     gattttttctt ttgaggcatg cacatctgga attaagggtc aactaattct cacatccctc 1957
106     taaaagtaaa ctactgttag gaacagcagt gttctcacag tgtggggcag ccgtccttct 2017
107     aatgaagaca atgatattga cactgtccct ctttggcagt tgcattagta actttgaaag 2077
108     gtatatgact gagcgtagca tacagggttaa cctgcagaaa cagtacttag gtaattgtag 2137
109     ggcgagggatt ataaatgaaa tttgcaaaat cacttagcag caactgaaga caattatcaa 2197
110     ccacgtggag aaaatcaaac cgagcagggc tgtgtgaaac atggttgtaa tatgcgactg 2257
111     cgaacactga actctacgcc actccacaaa tgatgttttc aggtgtcatg gactgttgcc 2317
112     acctgttatt catccagagt tcttaaagtt taaagttgca catgattgta taagcatgct 2377
113     ttctttgagt tttaaattat gtataaacat aagttgcatt tagaaatcaa gcataaatca 2437
114     cttcaactgc taaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 2479

```

115 &lt;210&gt; SEQ ID NO 2

116 &lt;211&gt; LENGTH: 350

117 &lt;212&gt; TYPE: PRT

118 &lt;213&gt; ORGANISM: Homo sapiens

119 &lt;400&gt; SEQUENCE: 2

```

120      Met Gln Arg Leu Gly Ala Thr Leu Leu Cys Leu Leu Leu Ala Ala Ala
121      1              5              10              15
122      Val Pro Thr Ala Pro Ala Pro Ala Pro Thr Ala Thr Ser Ala Pro Val
123      20              25              30
124      Lys Pro Gly Pro Ala Leu Ser Tyr Pro Gln Glu Glu Ala Thr Leu Asn
125      35              40              45
126      Glu Met Phe Arg Glu Val Glu Glu Leu Met Glu Asp Thr Gln His Lys
127      50              55              60
128      Leu Arg Ser Ala Val Glu Glu Met Glu Ala Glu Glu Ala Ala Ala Lys
129      65              70              75              80
130      Ala Ser Ser Glu Val Asn Leu Ala Asn Leu Pro Pro Ser Tyr His Asn
131      85              90              95
132      Glu Thr Asn Thr Asp Thr Asn Val Gly Asn Asn Thr Ile His Val His
133      100             105             110
134      Arg Glu Ile His Lys Ile Thr Asn Asn Gln Thr Gly Gln Met Val Phe
135      115             120             125
136      Ser Glu Thr Val Ile Thr Ser Val Gly Asp Glu Glu Gly Arg Arg Ser
137      130             135             140
138      His Glu Cys Ile Ile Asp Glu Asp Cys Gly Pro Ser Met Tyr Cys Gln
139      145             150             155             160
140      Phe Ala Ser Phe Gln Tyr Thr Cys Gln Pro Cys Arg Gly Gln Arg Met
141      165             170             175
142      Leu Cys Thr Arg Asp Ser Glu Cys Cys Gly Asp Gln Leu Cys Val Trp
143      180             185             190
144      Gly His Cys Thr Lys Met Ala Thr Arg Gly Ser Asn Gly Thr Ile Cys

```

PAGE: 4

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION US/09/009,802B**

DATE: 01/10/2000  
 TIME: 16:31:13

Input Set: I009802B.RAW

```

145          195          200          205
146    Asp Asn Gln Arg Asp Cys Gln Pro Gly Leu Cys Cys Ala Phe Gln Arg
147          210          215          220
148    Gly Leu Leu Phe Pro Val Cys Thr Pro Leu Pro Val Glu Gly Glu Leu
149          225          230          235          240
150    Cys His Asp Pro Ala Ser Arg Leu Leu Asp Leu Ile Thr Trp Glu Leu
151          245          250          255
152    Glu Pro Asp Gly Ala Leu Asp Arg Cys Pro Cys Ala Ser Gly Leu Leu
153          260          265          270
154    Cys Gln Pro His Ser His Ser Leu Val Tyr Val Cys Lys Pro Thr Phe
155          275          280          285
156    Val Gly Ser Arg Asp Gln Asp Gly Glu Ile Leu Leu Pro Arg Glu Val
157          290          295          300
158    Pro Asp Glu Tyr Glu Val Gly Ser Phe Met Glu Glu Val Arg Gln Glu
159          305          310          315          320
160    Leu Glu Asp Leu Glu Arg Ser Leu Thr Glu Glu Met Ala Leu Arg Glu
161          325          330          335
162    Pro Ala Ala Ala Ala Ala Ala Leu Leu Gly Arg Glu Glu Ile
163          340          345          350

```

```

164 <210> SEQ ID NO 3
165 <211> LENGTH: 1050
166 <212> TYPE: DNA
167 <213> ORGANISM: Homo sapiens
168 <220> FEATURE:
169 <221> NAME/KEY: CDS
170 <222> LOCATION: (1)..(1050)
171 <400> SEQUENCE: 3

```

```

172    atg cag cgg ctt ggg gcc acc ctg ctg tgc ctg ctg ctg gcg gcg gcg 48
173    Met Gln Arg Leu Gly Ala Thr Leu Leu Cys Leu Leu Leu Ala Ala
174          1          5          10          15
175    gtc ccc acg gcc ccc gcg ccc gct ccg acg gcg acc tcg gct cca gtc 96
176    Val Pro Thr Ala Pro Ala Pro Ala Pro Thr Ala Thr Ser Ala Pro Val
177          20          25          30
178    aag ccc ggc ccg gct ctc agc tac ccg cag gag gag gcc acc ctc aat 144
179    Lys Pro Gly Pro Ala Leu Ser Tyr Pro Gln Glu Glu Ala Thr Leu Asn
180          35          40          45
181    gag atg ttc cgc gag gtt gag gaa ctg atg gag gac acg cag cac aaa 192
182    Glu Met Phe Arg Glu Val Glu Glu Leu Met Glu Asp Thr Gln His Lys
183          50          55          60
184    ttg cgc agc gcg gtg gaa gag atg gag gca gaa gaa gct gct gct aaa 240
185    Leu Arg Ser Ala Val Glu Glu Met Glu Ala Glu Glu Ala Ala Lys
186          65          70          75          80
187    gca tca tca gaa gtg aac ctg gca aac tta cct ccc agc tat cac aat 288
188    Ala Ser Ser Glu Val Asn Leu Ala Asn Leu Pro Pro Ser Tyr His Asn
189          85          90          95
190    gag acc aac aca gac acg aac gtt gga aat aat acc atc cat gtg cac 336
191    Glu Thr Asn Thr Asp Thr Asn Val Gly Asn Asn Thr Ile His Val His
192          100          105          110
193    cga gaa att cac aag ata acc aac aac cag act gga caa atg gtc ttt 384
194    Arg Glu Ile His Lys Ile Thr Asn Asn Gln Thr Gly Gln Met Val Phe

```

PAGE: 5

# RAW SEQUENCE LISTING PATENT APPLICATION US/09/009,802B

DATE: 01/10/2000  
TIME: 16:31:13

Input Set: I009802B.RAW

```

195          115          120          125
196      tca gag aca gtt atc aca tct gtg gga gac gaa gaa ggc aga agg agc 432
197      Ser Glu Thr Val Ile Thr Ser Val Gly Asp Glu Glu Gly Arg Arg Ser
198          130          135          140
199      cac gag tgc atc atc gac gag gac tgt ggg ccc agc atg tac tgc cag 480
200      His Glu Cys Ile Ile Asp Glu Asp Cys Gly Pro Ser Met Tyr Cys Gln
201      145          150          155          160
202      ttt gcc agc ttc cag tac acc tgc cag cca tgc cgg ggc cag agg atg 528
203      Phe Ala Ser Phe Gln Tyr Thr Cys Gln Pro Cys Arg Gly Gln Arg Met
204          165          170          175
205      ctc tgc acc cgg gac agt gag tgc tgt gga gac cag ctg tgt gtc tgg 576
206      Leu Cys Thr Arg Asp Ser Glu Cys Cys Gly Asp Gln Leu Cys Val Trp
207          180          185          190
208      ggt cac tgc acc aaa atg gcc acc agg ggc agc aat ggg acc atc tgt 624
209      Gly His Cys Thr Lys Met Ala Thr Arg Gly Ser Asn Gly Thr Ile Cys
210          195          200          205
211      gac aac cag agg gac tgc cag ccg ggg ctg tgc tgt gcc ttc cag aga 672
212      Asp Asn Gln Arg Asp Cys Gln Pro Gly Leu Cys Cys Ala Phe Gln Arg
213          210          215          220
214      ggc ctg ctg ttc cct gtg tgc aca ccc ctg ccc gtg gag ggc gag ctt 720
215      Gly Leu Leu Phe Pro Val Cys Thr Pro Leu Pro Val Glu Gly Glu Leu
216          225          230          235          240
217      tgc cat gac ccc gcc agc cgg ctt ctg gac ctc atc acc tgg gag cta 768
218      Cys His Asp Pro Ala Ser Arg Leu Leu Asp Leu Ile Thr Trp Glu Leu
219          245          250          255
220      gag cct gat gga gcc ttg gac cga tgc cct tgt gcc agt ggc ctc ctc 816
221      Glu Pro Asp Gly Ala Leu Asp Arg Cys Pro Cys Ala Ser Gly Leu Leu
222          260          265          270
223      tgc cag ccc cac agc cac agc ctg gtg tat gtg tgc aag ccg acc ttc 864
224      Cys Gln Pro His Ser His Ser Leu Val Tyr Val Cys Lys Pro Thr Phe
225          275          280          285
226      gtg ggg agc cgt gac caa gat ggg gag atc ctg ctg ccc aga gag gtc 912
227      Val Gly Ser Arg Asp Gln Asp Gly Glu Ile Leu Leu Pro Arg Glu Val
228          290          295          300
229      ccc gat gag tat gaa gtt ggc agc ttc atg gag gag gtg cgc cag gag 960
230      Pro Asp Glu Tyr Glu Val Gly Ser Phe Met Glu Glu Val Arg Gln Glu
231          305          310          315          320
232      ctg gag gac ctg gag agg agc ctg act gaa gag atg gcg ctg agg gag 1008
233      Leu Glu Asp Leu Glu Arg Ser Leu Thr Glu Glu Met Ala Leu Arg Glu
234          325          330          335
235      cct gcg gct gcc gcc gct gca ctg ctg gga agg gaa gag att 1050
236      Pro Ala Ala Ala Ala Ala Leu Leu Gly Arg Glu Glu Ile
237          340          345          350
238      <210> SEQ ID NO 4
239      <211> LENGTH: 848
240      <212> TYPE: DNA
241      <213> ORGANISM: Homo sapiens
242      <220> FEATURE:
243      <221> NAME/KEY: CDS
244      <222> LOCATION: (125)..(796)

```

**Please Note:**

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Input Set: I009802B.RAW

Line	Error/Warning	Original Text
92	W "N" or "Xaa" used: Feature required	caatagaaat agctaattta tttccccang tgtgtgct
448	W "N" or "Xaa" used: Feature required	acccatttn attctagagt cnagaacgca aggatctc
593	W Invalid/Missing Amino Acid Numbering	